Dual Gauge Programs, with Applications to Quadratic Programming and The Minimum-Norm Problem

by

Robert M. Freund

Working Paper #1726-85 Revised August 1986

Abstract

A gauge function $f(\cdot)$ is a nonnegative convex function that is positively homogeneous and satisfies $f(0)=0$. Norms and pseudonorms are specific instances of a gauge function. This paper presents a gauge duality theory for a gauge program, which is the problem of minimizing the value of a gauge function $f(\cdot)$ over a convex set. The gauge dual program is also a gauge program, unlike the standard Lagrange dual. We present sufficient conditions on $f(\cdot)$ that ensure the existence of optimal solutions to the gauge program and its dual, with no duality gap. These sufficient conditions are relatively weak and are easy to verify, and are independent of any qualifications on the constraints. The theory is applied to a class of convex quadratic programs, and to the minimum l_p norm problem. The gauge dual program is shown to provide a smaller duality gap than the standard dual, in a certain sense discussed in the text.

Keywords: Gauge function, norm, quadratic program, Lagrange dual, duality.

Running Header: Dual Gauge Programs.

Introduction

A gauge function $f(\cdot):R^n \to R \cup \{+\infty\}$ is a nonnegative convex function that is positively homogeneous and satisfies $f(0)=0$. Norms and pseudonorms are specific instances of a gauge function. A gauge program is defined as an optimization problem of the form

$$\begin{aligned} P: \quad & \text{minimize } f(x) \\ & \text{subject to } Mx \geq b \end{aligned} \quad ,$$

where $f(\cdot)$ is a gauge function. Many problems in mathematical programming fall into this category, including strictly convex quadratic programming, linear programming, and the minimum norm problem on a polyhedron (see Luenberger [9]). Duality for programs similar to P have been studied by Eisenberg [2], whose work has most recently been generalized by Gwinner [7]. Glassey [6] has examined instances where explicit Lagrange duals of convex homogeneous programs like P can be stated, without reference to primal variables.

This paper presents a gauge duality theory for gauge programs that contrasts, but is related to, the Lagrange dual of P. In particular, the gauge dual D of P is also a gauge program, unlike its Lagrange dual. The gauge duality theory states that if z and v are feasible values of the primal and dual objective functions, then $z.v \geq 1$, with equality only if z and v are optimal values of P and D. This inequality is analagous to the weak duality relationship $z \geq v$ for the Lagrange dual. We present sufficient conditions on $f(\cdot)$ that ensure that optimal solutions to the dual gauge programs exist and that $z \cdot v=1$ for these solutions. These sufficient conditions are relatively weak and are easy to verify. They are independent of any qualification on the constraints of the gauge program, unlike the Slater condition for Lagrange duality, for example.

In the case of quadratic programming, the theory developed is applicable to a class of quadratic programs that is slightly broader than the class of strictly convex quadratic programs. The gauge dual is equivalent (by a monotone transformation) to a quadratic program different from the Lagrange dual.

Another application of the gauge duality theory is to the problem of minimizing the l_p norm of a vector over a polyhedron. The gauge dual is shown to provide a smaller duality gap (in a certain sense discussed in the text) than the standard dual, and hence provides a better lower bound on the primal objective value, for feasible values of the dual, than does the standard dual program.

A final application of the gauge duality theory is to linear programming, where the gauge dual is different (but equivalent to) the standard linear programming dual.

In order to lay the groundwork for the ensuing theory, Section 1 reviews basic polarity properties of gauge functions. Section 2 presents the gauge duality theory, which includes a weak duality theorem, and necessary conditions for a strong duality theorem to be valid. The duality theory of Section 2 is generalized to gauge programs with nonlinear constraints in Section 3. The theory of Section 2 is applied to selected mathematical programming problem in Section 4. This section first discusses convex quadratic programming, followed by a duality analysis of the minimum l_p norm problem, for which strictly convex quadratic programming is a special case. The discussion shows that the duality gap for the gauge dual is in a certain sense smaller than that of the standard dual. Section 4 concludes with an analysis of linear programming in the context of gauge duality theory.

1. Preliminaries

Let R denote the set of real numbers and let $\bar{R}=R\cup\{+\infty\}$. A function $f(\cdot):R^n\to\bar{R}$ is called a gauge if $f(\cdot)$ is convex, nonnegative, positively homogeneous (i.e., $f(\alpha x) = \alpha f(x)$ for $\alpha > 0$), and $f(0)=0$. Norms and pseudonorms are gauge functions. A gauge function need not be symmetric and can take on the value $+\infty$, unlike a pseudonorm or a norm.

An example of a gauge function is

$$f(x)=f(x_1,\ x_2,\ x_3) = \begin{cases} \|(x_1-x_3),(x_1-x_2)\|_2 & \text{if } 2x_1-x_2-x_3 = 0 \\ +\infty & \text{otherwise.} \end{cases}$$

Note that $f(x)$ is finite only on the plane $\{x\in R^3 | 2x_1-x_2-x_3 = 0\}$, and that $f(x) = 0$ for all $x = (\alpha,\alpha,\alpha)$. In this example, $f(.)$ is symmetric.

For any convex set $C \subset R^n$, the polar of C, denoted C^o, is defined by $C^o = \{y\in R^n | y^T x \leq 1 \text{ for all } x\in C\}$. C^o is a closed convex set containing the origin, and $C^{oo} = C$ if and only if C is a closed convex set containing the origin (see Rockafellar [12], p. 121). If $f(.)$ is a gauge function, and if C is defined by

$$C = \{x\in R^n |\ f(x) \leq 1\} \tag{1}$$

then $f(\cdot)$ can be represented by

$$f(x) = \inf\ \{u \geq 0 | x\in uC\}. \tag{2}$$

where by convention, we denote $\inf\ \phi = +\infty$. Furthermore, $f(.)$ is a closed function (i.e., all of the level sets of $f(.)$ are closed) if and only if C is a closed set. Also, for any closed convex set C that contains the origin, the function $f(.)$ defined by (2) is a closed gauge function, called the closed gauge function corresponding to C.

For any gauge function $f(.)$, define its polar function $f^{\circ}(.)$ by

$$f^{\circ}(y) = \inf \{v \geq 0 \mid y^T x \leq v f(x) \text{ for all } x\}. \quad (3)$$

Then $f^{\circ}(.)$ is a closed gauge. If $C = \{x \mid f(x) \leq 1\}$, then $C^{\circ} = \{y \mid f^{\circ}(y) \leq 1\}$, whereby $f^{\circ}(.)$ is closed, since C° is closed. Furthermore, if $f(\cdot)$ is closed, then $f^{\circ\circ}(x) = f(x)$, because $C^{\circ\circ} = C$. The following summarizes the above statements:

Remark 1 (see Rockafellar [12], p. 129). The polarity operation $f(.) \to f^{\circ}(.)$ induces a one-to-one symmetric correspondence in the class of all closed gauges on R^n. Two closed convex sets containing the origin are polar to each other if and only if their gauge functions are polar to each other. [X]

We also have:

Remark 2. If $f(\cdot)$ is a closed gauge function, then $f(\cdot)$ and $f^{\circ}(\cdot)$ can be written as:

$$f(x) = \sup_{y \in C^{\circ}} y^T x \quad \text{and} \quad f^{\circ}(y) = \sup_{x \in C} y^T x \quad .$$

In the case when $f(x) = \|x\|_p$, the l_p norm, and $1 \leq p \leq \infty$, then $f^{\circ}(y) = \|x\|_q$, where $1/p + 1/q = 1$, and the Holder inequality states that $y^T x \leq \|x\|_p \|y\|_q = f(x) f^{\circ}(y)$. The following generalization of the Holder inequality can be stated as:

Remark 3. If $f(\cdot)$ is a gauge function, then $y^T x \leq f(x) f^{\circ}(y)$ provided $f(x)$ and $f^{\circ}(y)$ are both finite or that $\{f(x), f^{\circ}(y)\} \neq \{0, \infty\}$.

For a given function $f(\cdot): R^n \to \bar{R}$, its conjugate $f^*(\cdot)$ is defined by $f^*(y) = \sup_x \{y^T x - f(x)\}$. If $f(\cdot)$ is a gauge function, then it is straightforward and demonstrate that

$$f^*(y) = \begin{cases} 0 \text{ if } f^\circ(y) \leq 1 \\ +\infty \text{ if } f^\circ(y) > 1 \end{cases} .$$

If $g(x)$ is a gauge function defined by $x \in R^n$ and $w \in R^m$, then the function $f(x,w):R^{n+m} \to \bar{R}$ defined by $f(x,w) = g(x)$ is a gauge function, and

$$f^\circ(y,z) = \begin{cases} g^\circ(y) & \text{if } z = 0 \\ +\infty & \text{if } z \neq 0 \end{cases} ,$$

where $y \in R^n$ and $z \in R^m$.

2. Dual Gauge Programs

Consider the following nonlinear program:

$$\text{P:} \quad \text{minimize} \quad z = f(x)$$
$$\text{subject to} \quad Mx \geq b$$

where $f(\cdot)$ is a closed gauge function. The Lagrange dual of P is formulated as

$$\underset{\lambda \geq 0}{\text{supremum}} \{\underset{x}{\text{infimum}} \ \{f(x) - \lambda^T(Mx-b)\}\},$$

which can be simplified to

$$\underset{\lambda \geq 0}{\text{supremum}} \{b^T\lambda - f^*(M^T\lambda)\},$$

or

$$\text{LD:} \quad \text{maximize} \quad v = b^T\lambda$$
$$\text{subject to} \quad f^\circ(M^T\lambda) \leq 1$$
$$\lambda \geq 0 \quad .$$

Dual programs for classes of programs that include P have been developed by Eisenberg [2] and Gwinner [7]. Glassey [6] has shown how to construct explicit duals (with no primal variables) for such problems. All three authors work with a primal problem for which the objective function $f(\cdot)$ is convex and positively homogeneous ($f(\cdot)$ need not be nonnegative, as in our primal, but is restricted to be finite-valued). When applied to a gauge program P, however, the dual programs that each author develops is the program LD above.

If we exchange the objective function with the polar gauge constraint in LD, we obtain a dual gauge program:

$$\text{D:} \quad \text{minimize} \quad v = f^\circ(M^T\lambda)$$
$$\text{subject to} \quad b^T\lambda = 1$$
$$\lambda \geq 0 \quad .$$

Together, the pair P and D constitute dual gauge programs, Note that both the primal (P) and dual (D) are minimization problems, and their objective functions are polar gauge functions. The dual variables λ are restricted to be nonnegative and correspond to primal constraints. The constraint matrix M in the primal appears in the dual in the objective function, and the right-hand side (RHS) of the primal appears in the equality constraint in the dual.

The definition of the gauge dual program D can be extended to other types of linear constraints in the standard way. If the i^{th} constraint of P is $M_i x = (\leq) b_i$, then in the dual D, we require the i^{th} variable λ_i to be unrestricted in sign (less than or equal to zero.)

Note that the dual of the dual is the primal. To see this, write the dual in the format:

$$\begin{array}{lrcll} \text{minimize} & v & = & g(y,\lambda) & \\ \text{subject to:} & y-M^T\lambda & = & 0 & \text{(i)} \quad (x) \\ & b^T\lambda & = & 1 & \text{(ii)} \quad (t) \\ & \lambda & \geq & 0 & \text{(iii)} \quad (s). \end{array}$$

where $g(y,\lambda) = f^{\circ}(y)$. Then if we associate the variables x, t, and s with constraints (i), (ii), and (iii), the dual of this program is:

$$\begin{array}{ll} \text{minimize} & z = g^{\circ}(x, bt-Mx+s) \\ \text{subject to:} & t = 1 \\ & s \geq 0. \end{array}$$

However, $g^{o}(x,w) = f^{oo}(x) = f(x)$ when $w=0$, and $g^{o}(x,w) = +\infty$ if $w \neq 0$. Thus the last program can be written as

$$\begin{aligned} &\text{minimize} && z = f(x) \\ &\text{subject to:} && Mx-s = b \\ & && s \geq 0 \end{aligned} \quad ,$$

which is precisely the primal P.

The vector of variables x (λ) is said to be <u>feasible</u> for P (D) if x (λ) satisfy the linear constraints, i.e. $Mx \geq b$ ($b^T\lambda=1$, $\lambda \geq 0$), otherwise x (λ) is infeasible. If x (λ) is feasible but $f(x) = +\infty$ ($f^{o}(M^T\lambda) = +\infty$), then x (λ) is <u>essentially infeasible</u>. If $x(\lambda)$ is feasible and $f(x) < +\infty$ ($f^{o}(M^T\lambda) < +\infty$), then x (λ) is <u>strongly feasible</u> If P (D) has no strongly feasible solution, P (D) is an <u>essentially infeasible program</u>; otherwise P (D) is a <u>strongly feasible program</u>.

We have the following preliminary duality result for dual gauge programs.

<u>Theorem 1</u>. Let z^* and v^* be optimal values of (P) and (D), respectively. Then:

(i) If x and λ are strongly feasible for P and D, with objective values z and v, respectively, then $zv \geq 1$, and hence $z^*v^* \geq 1$.

(ii) If $z^* = 0$, then D is essentially infeasible, i.e., $v^* = +\infty$.

(iii) If $v^* = 0$, then P is essentially infeasible, i.e., $z^* = +\infty$.

(iv) If $\bar{x}$ and $\bar{\lambda}$ are feasible solutions for P and D with objective values $\bar{z}$ and $\bar{v}$, respectively, and $\bar{z}\bar{v} = 1$, then $\bar{x}$ and $\bar{\lambda}$ are optimal solutions of P and D, respectively.

<u>PROOF</u>: If x and λ are strongly feasible for P and D, then we have $1 = \lambda^T b \leq \lambda^T Mx \leq f^{o}(M^T\lambda)\, f(x) = zv$, by Remark 3. This result shows (i), and (ii) and (iii) follow by contradiction from (i). (iv) follows

from (i) by noting that $1/\bar{v}$ is a lower bound on z^*, and is achieved by $\bar{z}$. [X]

Assertion (i) corresponds to the standard duality result that the value of the max program is less than or equal to the value of the min program, and assertion (iv) corresponds to the result that if max equals min, then both are optimal. Assertions (ii) and (iii) correspond to unbounded cases in the standard theory. Because $f(.)$ is a gauge, $f(x) \geq 0$ for any x, whereby $z^* \geq 0$. If $z^* = 0$, the program P has achieved its absolute lower limit, in the same way that a standard program would have a value of $-\infty$, and hence the dual program is infeasible, i.e., $v^* = +\infty$.

In order to prove a strong duality theorem for the dual gauge programs P and D that asserts that $z^*v^*=1$ and that P and D both attain their optimal values, it is necessary to impose some qualifications on the function $f(.)$. We proceed as follows.

A convex set $S \subset R^n$ satisfies the projection property if all projections of S are closed convex sets, i.e., if for any linear transformation $A:R^n \rightarrow R^m$, $\{z \in R^m | z=Ax \text{ for some } x \in S\}$ is a closed convex set. A gauge function $f(\cdot)$ satisfies the projection qualification if both C and C^o satisfy the projection property, where C is given by $C = \{x \in R^n | f(x) \leq 1\}$. For notational convenience, $f(\cdot)$ and C will be assumed to be related by relations (1) and (2) of the previous section, for the remainder of this paper.

Note that by definition, $f(\cdot)$ satisfies the projection qualification if and only if $f^o(\cdot)$ satisfies the projection qualification. If C and C^o are convex and compact, then $f(\cdot)$ satisfies the projection qualification, and hence $f(x) = \|x\|_p$ satisfies the projection qualification for $1 \leq p \leq \infty$. Also note that if C is a polyhedron (bounded or not), then C^o is a polyhedron and $f(\cdot)$ and $f^o(\cdot)$ satisfy the projection qualification. Finally, note that if $g(\cdot)$ is a gauge function that satisfies the projection qualification, and $f(x,w) = g(x)$ then $f(\cdot,\cdot)$ satisfies the projection qualification.

The projection qualification allows us to prove the following results which will be useful in proving the strong duality theorem.

Remark 4. If $f(.)$ satisfies the projection qualification and if $f(\bar{x})$ is finite, then $f(\bar{x}) = \bar{y}^T\bar{x}$ for some $\bar{y} \in C^o$.

PROOF: We have $f(\bar{x}) = \sup_{y \in C^o} y^T\bar{x} = \sup \{u \mid u = y^T\bar{x} \text{ for some } y \in C^o\}$.

However, $\{u \mid u = y^T\bar{x} \text{ for some } y \in C^o\}$ is a nonempty closed convex set, i.e. a closed interval, by the projection qualification. If $f(\bar{x})$ is finite, then $f(\bar{x}) = \bar{y}^T\bar{x}$ for some $\bar{y} \in C^o$. [X]

Lemma 1. If $f(.)$ satisfies the projection qualification, and X is a polyhedron such that $\mu C \cap X = \phi$ for a given $\mu > 0$, then there exists a vector $y \in R^n$ such that for all $x \in \mu C$, $y^Tx < 1$, and for all $x \in X$, $y^Tx \geq 1$.

PROOF: The sets μC and X are convex and have no points in common. There thus exists a hyperplane that separates them. Because $f(.)$ satisfies the projection qualification, all projections of μC are closed. Furthermore, X is polyhedron. Consequently, according to remark 5 in the appendix, there exists a hyperplane that separates μC from X and does not meet μC. Therefore, there exists $(y,\alpha) \in (R^n,R)$ such that $y^Tx < \alpha$ for all $x \in \mu C$, and $y^Tx \geq \alpha$ for all $x \in X$. Since $x=0 \in \mu C$, α must be positive, whereby by scaling we can presume it is equal to 1, completing the proof. [X]

We can now prove:

Theorem 2. Assume that f(.) satisfies the projection qualification, and let z^* and v^* be the optimal values of P and D, respectively. Then:

(i) If P and D are both strongly feasible, then $z^*v^* = 1$; and the optimal values of P and D are achieved for some x^* and λ^*.

(ii) P is essentially infeasible (i.e. $z^*=\infty$) if and only if $v^*=0$; and $v^*=0$ is achieved for some feasible λ^* if P is infeasible.

(iii) D is essentially infeasible (i.e. $v^*=\infty$) if and only if $z^*=0$; and $z^*=0$ is achieved for some feasible x^* if D is infeasible.

PROOF: (i) Let $X = \{x \in R^n | Mx \geq b\}$. If P and D are strongly feasible, then $0 < z^* < +\infty$ and $0 < v^* < +\infty$. We must now show that for some $x \in X$, $f(x) = 1/v^*$. Assume the contrary. Then $(1/v^*)C \cap X = \phi$, and from Lemma 1, there exists $y \in R^n$ that satisfies $y^T x < 1$ for all $x \in (1/v^*)C$, and $y^T x \geq 1$ for all feasible x. But since C satisfies the projection property, so does $(1/v^*)C$, and therefore $\{y^T x | x \in (1/v^*)C\}$ is a closed convex set, i.e. a closed interval $[c,d]$ or $(-\infty,d]$, and $d < 1$. Now since $y^T x \geq 1$ for all x that satisfy $Mx \geq b$, there exists $\lambda \geq 0$ with $y = M^T\lambda$, $\lambda^T b \geq 1$. Now $\bar{\lambda} = \lambda/\lambda^T b$ is feasible for D, and $f^\circ(M^T\bar{\lambda}) = (1/\lambda^T b)\, f^\circ(y) \leq dv^* < v^*$, a contradiction. Thus P achieves its minimum at some feasible point x^*, and $z^* = f(x^*) = 1/v^*$.

A parallel argument shows that D achieves its minimum at some feasible point λ^*, establishing (i). Regarding (ii), the "if" part of the statement has been shown in Theorem 1. For the "only if" part, assume that P is essentially infeasible. Then for any finite $\mu > 0$, $\mu C \cap X = \phi$, whereby there exists a vector $y \in R^n$ as described in

Lemma 1. Let λ and $\bar{\lambda}$ be constucted as above. Then $\bar{\lambda}$ is feasible for D and $f^{\circ}(M^T\bar{\lambda}) \leq 1/u$, whereby $v^* \leq 1/u$ for any $u > 0$, i.e. $v^*=0$.

If P is infeasible, then by a theorem of the alternative, there exists $\bar{\lambda} \geq 0$ with $M^T\bar{\lambda} = 0$ and $b^T\bar{\lambda} = 1$. Furthermore $f^{\circ}(M^T\bar{\lambda}) = 0 = v^*$. This completes the proof of (ii).

The proof of (iii) parallels that of (ii). [X]

Theorem 2 thus provides a rather weak qualification on $f(\cdot)$ that is sufficient to guarantee the existence of primal and dual optimal solutions, namely that all projections of C and C° be closed. Note that the projection qualification makes no reference to the constraints of P, in contrast to more typical sufficient conditions such as the Slater condition, which is used directly in Glassey [6] and indirectly in Eisenberg [2] and Gwinner [7]. As the proof of Theorem 2 indicates, a constraint qualification is not needed, because the feasible region is polyhedral. Also note that $f(\cdot)$ need not be differentiable, and can take on the value $+\infty$ in the feasible region of P. The proof of Theorem 2 is based essentially on arguments stemming from an "open separation" theorem (see the appendix), which states that if uC and X are disjoint for some $u > 0$, then uC can be openly separated from X. The projection qualification is sufficient to guarantee the open separation.

Although part (i) of Theorem 2 asserts that the projection qualification is sufficient in order for a gauge program to attain its optimum z^* where $0 < z^* < \infty$, part (ii) makes no such assertion when $z^*=0$. The rather stronger condition that the dual is infeasible is shown to be sufficient in this case. To see that the projection qualification is not sufficient to guarantee the existence of an optimal solution when $z^*=0$, consider the following example. Let $C = \{(x_1,x_2)\in R^2 | x_2 \geq x_1^2/2\}$ and let $f(\cdot)$ be defined as in (2). Thus we have

$$f(x_1,x_2) = \begin{cases} x_1^2/(2x_2) & x_2 > 0 \\ 0 & x_1=0,\ x_2=0 \\ +\infty & x_1\neq 0,\ x_2=0 \\ +\infty & x_2 < 0 \end{cases}$$

It is straightforward to derive $C^o = \{(y_1,y_2)\in R^2 \mid y_2 \leq -y_1^2/2\}$, and to verify that both C and C^o satisfy the projection qualification.

Let the feasible region X of P be defined by $X = \{x\in R^2 \mid Mx \geq b\}$, where

$$M = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{, and } b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

By choosing $\bar{x}_1=1$ and $\bar{x}_2$ sufficiently large, $M\bar{x} \geq b$, and $f(\bar{x}_1,\bar{x}_2) = \bar{x}_1^2/(2\bar{x}_2) = 1/(2\bar{x}_2)$, which goes to zero as $\bar{x}_2$ goes to infinity. Thus $z^*=0$. However, there is no feasible pair (x_1,x_2) for which $f(x_1,x_2)=0$.

An alternate sufficient condition for strong duality in P and D in the spirit of the Slater condition is given below. We proceed as follows. Given dual gauge programs P and D, define $X = \{x\in R^n \mid Ax \geq b\}$, and $Y = \{y\in R^n \mid y=A^T\lambda,\ \lambda \geq 0,\ b^T\lambda \geq 1\}$. Let $\tilde{\tilde{C}} = \{x\in R^n \mid f(x) < \infty\}$ and $\tilde{\tilde{C}}^o = \{y\in R^n \mid f^o(y) < \infty\}$. We have

<u>Lemma 2</u>. The dual gauge programs P and D each attain their optima z^* and v^* and $z^*v^* = 1$, if $(\text{rel int } X) \cap (\text{rel int } \tilde{\tilde{C}}) \neq \Phi$, and $(\text{rel int } Y) \cap (\text{rel int } \tilde{\tilde{C}}^o) \neq \Phi$.

Note that the condition on the intersection of relative interiors is precisely Fenchel's sufficient condition for strong duality, which has been shown to be equivalent to the Slater condition for the Lagrange dual, see Magnanti [10]. Unlike the projection qualification, this condition can be rather cumbersome to verify in practice.

We will not prove Lemma 2 here. Its proof follows as an immediate consequence of Theorem 2A of the next section, which is a restatement of Lemma 2 for a gauge program with nonlinear constraints. As the proof there indicates, the intersection condition of Lemma 2 guarantees strong duality by guaranteeing proper separation of z^*C and X if z^* is not attained. The projection qualification, on the other hand, gives us strong duality by guaranteeing open separation of z^*C and X if z^* is not attained. In this sense, the projection qualification is a stronger qualification, because open separation is a stronger separation than proper separation. Nevertheless, the projection qualification is indeed a "sufficiently" weak condition so as to be valid for all of the applications in Section 4 of this paper.

3. An Approach to Dual Gauge Programs with Nonlinear Constraints

In this section, we show how the gauge duality theory developed in Section 2 for gauge programs (with linear constraints) can conceptually be extended to include nonlinear constraints. Consider the nonlinear gauge program

NGP: minimize $z = f(x)$

subject to $x \in X$

where $f(\cdot)$ is a closed gauge function, and X is a closed convex set, not necessarily a polyhedron.

Corresponding to the gauge function $f(\cdot)$ is its polar function $f^o(\cdot)$. In order to develop a nonlinear gauge dual of NGP, we also use a duality correspondence for the set X. For a given closed convex set X, define X' by the relation

$$X' = \{y \in R^n \mid y^T x \geq 1 \text{ for all } x \in X\}.$$

Following McLinden [11], we will refer to X' as the antipolar of X, although this nomenclature is not universal. (In a more restrictive context, X' is the blocker of X in Fulkerson [5]. In Ruys [13], X' is the upper dual set of X.) We define the nonlinear gauge dual of NGP to be the program

NGD: minimize $v = f^o(y)$

subject to $y \in X'$.

In the dual, the objective function is the polar of the primal gauge function, and the dual objective function is the antipolar of the primal feasible region. In order to characterize when the dual of the dual is the primal, we need to introduce some additional definitions. For a given nonempty convex set $X \subset R^n$, a vector $r \in R^n$ is called a ray of X if for every $x \in X$, $x + \theta r \in X$ for all $\theta \geq 0$. X is a ray-like set if every element x of X is also a ray of X. (McLinden [11] calls such an X an <u>antipolar</u> set, Ruys [13] calls X <u>auerole-reflexive</u>).

<u>Lemma 3</u> (see also McLinden [11], p. 176). For any set $X \subset R^n$, its antipolar X' is a raylike set. If X is closed, convex, raylike, and does not contain the origin, then X" = X.

<u>PROOF</u>: If $X=\phi$, then $X'=R^n$, which is raylike, and X"=X. If $X\neq\phi$, then X' is the intersection of a family of closed halfspaces, and so is closed and convex. If $y \in X'$, then $\theta y \in X'$ for all $\theta \geq 1$, and so X' is raylike.

We now must show that if X is a nonempty closed, convex, raylike set that does not contain the origin, then X"=X. Let $x \in X$. Then $y^T x \geq 1$ for all $y \in X'$, whereby $x \in X''$, and so $X \subset X''$. Suppose $X \neq X''$. Then there exists an element z of X" that is not contained in X. Because X is closed and convex, there exists a hyperplane that strictly separates {z} from X, and so there exists $(y,\alpha) \in (R^n, R)$, with the property that $y^T x > \alpha$ for all $x \in X$, and $y^T z < \alpha$. If $\alpha > 0$, then by rescaling we can assume that $\alpha=1$. This being the case, $y \in X'$, and so $y^T z \geq 1=\alpha$, a contradiction. Thus $\alpha \leq 0$. Because X is raylike, $y^T(\theta x) > \alpha$ for all $x \in X$ and all $\theta \geq 1$, and hence $y^T x \geq 0$ for all $x \in X$. Also $y^T z < \alpha \leq 0$. Because $0 \notin X$ and X is nonempty, X' is nonempty. Let $\bar{y}$ be any element of X'. Then, $\bar{y}^T x \geq 1$ for every $x \in X$. Because $y^T x \geq 0$ for every $x \in X$, $(\bar{y} + \theta y)^T x \geq 1$ for any $\theta \geq 0$, for every $x \in X$. Therefore $(\bar{y} + \theta y) \in X'$ for every $\theta \geq 0$. This in turn implies that $(\bar{y} + \theta y)^T z \geq 1$ for all $\theta \geq 0$, and so $y^T z \geq 0$, which contradicts $y^T z < \alpha \leq 0$. [X]

Lemma 3 implies that the dual of NGD is precisely NGP whenever $f(\cdot)$ is closed, and X is a closed, convex, raylike set that does not contain the origin. Of course, if X does contain the origin, then

$z^*=0$, $X'=\phi$, and $v^*=\infty$, where z^* and v^* are the optimal values of the primal and dual, respectively.

The following result is analogous to Theorem 1:

Theorem 1A (Weak Duality). Let z^* and v^* be optimal values for NGP and NGD, respectively. Then

(i) If x and y are strongly feasible for NGP and NGD, with objective values z and v, respectively, the $zv \geq 1$, and hence $z^*v^* \geq 1$.

(ii) If $z^*=0$, then NGD is essentially infeasible, i.e. $v^*=+\infty$.

(iii) If $v^*=0$, then NGP is essentially infeasible, i.e. $z^*=+\infty$

(iv) If $\bar{x}$ and $\bar{y}$ are strongly feasible for NGP and NGD with objective values $\bar{z}$ and $\bar{v}$, respectively, and $\bar{z}\bar{v} = 1$, then $\bar{x}$ and $\bar{y}$ are optimal solutions of NGP and NGD, respectively.

PROOF: If x and y are strongly feasible for NGP and NGD, then $1 \leq y^Tx \leq f^\circ(y)\, f(x) = zv$, by Remark 3. This result shows (i), and (ii), (iii), and (iv) follow from (i). [X]

To see how to obtain the linearly constrained problems P and D from NGP and NGD, let P be as given. Define $X = \{x \in R^n | Mx \geq b\}$ and $\bar{X} = \{x \in R^n | x \in \theta X \text{ for some } \theta \geq 1\} = \{x \in R^n | Mx \geq b\theta \text{ for some } \theta \geq 1\}$. Then $\bar{X}$ is a raylike set that contains X. Furthermore, $\bar{X}' = \{y \in R^n | y=M^T\lambda \text{ for some } \lambda \geq 0 \text{ satisfying } b^T\lambda \geq 1\}$, and define $Y = \{y \in R^n | y=M^T\lambda \text{ for some } \lambda \geq 0 \text{ satisfying } b^T\lambda = 1\}$, and note that $\bar{X}'$ is raylike and contains Y. The linearly constrained gauge program P is equivalent to the program

$\bar{P}$: minimize $f(x)$

subject $x \in \bar{X}$

because even though the feasible region of $\bar{P}$ contains the feasible region of P, every point $\bar{x}\in\bar{X}$ has at least as large an objective function value as a corresponding point $x\in X$. The nonlinear gauge dual of $\bar{P}$ is

$$
\begin{aligned}
\bar{D}: \quad & \text{minimize } f^\circ(M^T\lambda) \\
& \text{subject to } b^T\lambda \geq 1 \\
& \qquad\qquad\quad \lambda \geq 0 \quad .
\end{aligned}
$$

However, because $f^\circ(\cdot)$ is homogeneous, we can restrict our attention to those $\lambda \geq 0$ for which $b^T\lambda = 1$, obtaining the equivalent program:

$$
\begin{aligned}
D: \quad & \text{minimize } f^\circ(M^T\lambda) \\
& \text{subject to } b^T\lambda = 1 \\
& \qquad\qquad\quad \lambda \geq 0 \quad ,
\end{aligned}
$$

which is the linear gauge dual.

We have the following strong duality result for the nonlinear case which is an extension of Lemma 2. Let $\tilde{\tilde{C}} = \{x\in R^n | f(x) < \infty\}$ and $\tilde{\tilde{C}}^\circ = \{y\in R^n | f^\circ(y) < \infty\}$.

<u>Theorem 2A</u>. Given dual gauge programs NGP and NGD, where $f(\cdot)$ is closed and X is closed, convex, and raylike, let z^* and v^* be optimal values of NGP and NGD, respectively. If (rel int X) $\cap$ (rel int $\tilde{\tilde{C}}$) $\neq \phi$ and (rel int X') $\cap$ (rel int $\tilde{\tilde{C}}^\circ$) $\neq \phi$, then $z^*v^* = 1$, and each program attains its optimum.

<u>PROOF</u>: Let $x^\circ \in$ (rel int X) $\cap$ (rel int $\tilde{\tilde{C}}$) and $y^\circ \in$ (rel int X') $\cap$ (rel int $\tilde{\tilde{C}}^\circ$). Then both NGP and NGD are strongly feasible, and by Theorem 1A, $0 < z^* < \infty$. Suppose that z^* is not attained by any feasible $x\in X$. Then $z^*C \cap X = \phi$, and so z^*C and X can be properly separated by a hyperplane H. Thus, there exists $\bar{y}\in R^n$ and $\alpha\in R$ such that $\bar{y}^Tx \geq \alpha$ for all $x\in X$, and $\bar{y}^Tx \leq \alpha$ for all $x\in z^*C$. If $\alpha > 0$, then we can presume $\alpha = 1$ by rescaling if necessary. Then $\bar{y}\in X'$

and $v^* \leq f^\circ(\bar{y}) = \sup_{x \in C} \bar{y}^T x = (1/z^*) \sup_{x \in z^*C} \bar{y}^T x \leq \alpha/z^* = 1/z^*$.

However, since $v^* z^* \geq 1$, we have $z^* v^* = 1$ and $\bar{y}$ is optimal for NGD. If $\alpha \leq 0$, then $\alpha = 0$, because $0 \in z^*C$. Then $\bar{y}^t x^\circ \geq 0$ because $x^\circ \in X$. Also, because $x^\circ \in$ rel int $\tilde{\tilde{C}}$, $(z^*/f(x^\circ))x^\circ \in z^*C$, whereby $\bar{y}^T x^\circ \leq 0$. Thus $\bar{y}^T x^\circ = 0$. Because $x^\circ \in$ rel int X, for every $x \in X$, there exists $\delta > 1$ such that $\delta x^\circ + (1-\delta)x \in X$. Thus $\bar{y}^T(\delta x^\circ + (1-\delta)x) \geq 0$, which implies $\bar{y}^T x \leq 0$. This in turn means $\bar{y}^T x = 0$ for all $x \in X$. Similarly, we can demonstrate that $\bar{y}^T x = 0$ for all $x \in z^*C$, and hence for all $x \in \tilde{\tilde{C}}$. Thus H does not properly separate X from z^*C. This contradiction ensures that $\alpha > 0$, and so v^* is attained in the dual and $z^* v^* = 1$. If z^* is attained in the primal, then the above proof is still valid, so long as z^*C and X can be properly separated by a hyperplane H. If z^*C and X cannot be properly separated, then by Theorem 6 of the appendix, there exists $\bar{x} \in$ (rel int X)$\cap$(rel int z^*C). Because $\bar{x} \in$ rel int z^*C and $0 \in z^*C$, there exists $\delta > 1$ such that $\delta \bar{x} + (1-\delta)0 \in z^*C$, whereby $\bar{x} \in (z^*/\delta)C$, and so $f(\bar{x}) < z^*$, a contradiction. Thus z^*C and X can be properly separated, and so v^* is attained in the dual, and $z^*v^*=1$.

A parallel argument establishes that z^* is attained in the primal. [X]

The nonlinear gauge duality theory parallels the gauge duality theory for linear constraints. It is only natural then to examine if there is a parallel duality construction that extends the Lagrange-type dual LD to handle nonlinear constraints. If X is closed, convex, and does not contain the origin, NGP can be written as

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & y^T x \geq g(y) \text{ for all } y \in \text{cone } X', \end{array}$$

where cone $X' = \{y \in R^n | y = \alpha w$ for some $\alpha \geq 0$ and $w \in X'\}$ and $g(y) = \max \{\alpha \geq 0 | y = \alpha w$ for some $w \in X'\}$. The above program is in a suitable format so that Gwinner's dual [7] can be constructed, which is:

$$\begin{array}{lll} \text{GwD:} & \text{maximize} & g(y) \\ & \text{subject to} & y^T x \leq f(x) \text{ for all } x \\ & & y \in \text{cone } X' \end{array}$$

Because $y^T x \leq f(x)$ for all x if and only if $f^\circ(y) \leq 1$, program GwD can be transformed into

$$\begin{array}{lll} \text{NLD:} & \underset{y,\alpha}{\text{maximize}} & \alpha \\ & \text{subject to} & f^\circ(\alpha y) \leq 1 \\ & & y \in X' \end{array},$$

which we define as the program NLD. To see that NLD and GwD are identical, notice that for any feasible solution (y,α) to NLD, $y' = (\alpha/g(y))y$ is feasible for GwD with identical objective value. For every feasible solution y' of GwD, $y' = \beta w$ for some $\beta \geq 0$ and $w \in X'$, and $y = (1/d(w))y'$, $\alpha = \beta d(w)$ is feasible for NLD.

The dual nonlinear gauge programs NGP and NGD make up a neat theory in terms of polar functions and antipolar sets. However, the explicit dual variables y do not directly correspond to primal constraints (though they do indirectly), and there is no formal mention of constraints *per se* in either the primal or the dual. One special case of the general nonlinear theory is of course the linear theory, instances of which will be seen in the next section. An open issue regarding nonlinear gauge duality is under what circumstances can the nonlinear gauge dual program be written explicitly in terms of a finite number of constraints whose variables include a multiplier for every primal constraint?

4. Applications

In this section we explore a number of mathematical programming models that correspond to a gauge program, including convex quadratic programming, problems involving the l_p norm, and linear programming. Many of the applications will involve programs of the form:

$$\begin{array}{ll} \underset{x}{\text{minimize}} & f(Nx+d) \\ \text{subject to} & Ax \geq b \end{array}$$

where $f(\cdot)$ is a gauge. Note that this format does not conform to that of P. However, it is equivalent to:

$$\begin{array}{lll} \underset{x,s}{\text{minimize}} & g(x,s) & \\ \text{subject to} & Ax & \geq b \\ & -Nx + Is & = d \end{array}$$

where $g(x,s) = f(s)$. Furthermore, $f(\cdot)$ satisfies the projection qualification if and only if $g(.,.)$ does. The gauge dual of this program is:

$$\begin{array}{ll} \underset{\lambda,\mu}{\text{minimize}} & g^{\circ}(A^T\lambda - N^T\mu, \mu) \\ \text{subject to} & b^T\lambda + d^T\mu = 1 \\ & \lambda \geq 0 \end{array}$$

But $g^{\circ}(y,t) = f^{\circ}(t)$ when $y=0$, and $g^{\circ}(y,t) = +\infty$ for $y \neq 0$; thus this last program becomes

$$\begin{array}{ll} \underset{\lambda,\mu}{\text{minimize}} & f^{\circ}(\mu) \\ \text{subject to} & A^T\lambda - N^T\mu = 0 \\ & b^T\lambda + d^T\mu = 1 \\ & \lambda \geq 0 \end{array} .$$

A. Convex Quadratic Programming

The standard convex quadratic program is given by

$$\text{QP:} \qquad \begin{array}{ll} \text{minimize} & (1/2)x^TQx + q^Tx \\ \text{subject to} & Ax \geq b \end{array}$$

where Q is a symmetric positive semi-definite matrix. Furthermore, the matrix Q can be factored into the form $Q = M^TM$ for some square matrix M. If M is nonsingular (i.e., Q is positive definite), or if q lies in the row space of M, then $q = M^Ts$ for some vector $s \in R^n$, and QP can be written as

$$\begin{array}{ll} \text{minimize} & (1/2)x^TM^TMx + s^TMx \\ \text{subject to} & Ax \geq b \end{array} \quad ,$$

which is equivalent to the gauge program

$$\text{GP:} \qquad \begin{array}{ll} \text{minimize} & \|Mx+s\|_2 \\ \text{subject to} & Ax \geq b \end{array} \quad ,$$

where $f(\cdot) = \|\cdot\|_2$, and so $f(\cdot)$ satisfies the projection qualification. Note that QP and GP are equivalent in that their constraints are identical and their objective functions differ by a strictly monotone transformation. In examining the duality properties of QP and GP, we will first study the case where Q is positive definite, followed by the case when Q is positive semi-definite.

Q is Positive Definite

When Q is positive definite the Lagrange dual of QP is

$$\begin{array}{ll} \underset{\lambda}{\text{maximize}} & b^T\lambda - (1/2)(\lambda^TA-q^T)Q^{-1}(A^T\lambda-q) \\ \text{subject to} & \lambda \geq 0 \end{array} \qquad \text{(LQP)},$$

see Dorn [1]. The guage dual of GP, on the other hand, is

$$\begin{array}{ll} \underset{\lambda}{\text{minimize}} & \|(M^T)^{-1}A^T\lambda\|_2 \\ \text{subject to} & (b^T+q^TQ^{-1}A^T)\lambda = 1 \\ & \lambda \geq 0 \end{array} \qquad \text{(GD)},$$

which is equivalent to the quadratic program

$$\begin{array}{ll} \text{minimize} & \lambda^T A Q^{-1} A^T \lambda \\ \text{subject to} & (b^T + q^T Q^{-1} A^T)\lambda = 1 \\ & \lambda \geq 0 \end{array} \qquad (\overline{GD}) \ .$$

Note that both LQP and $\overline{GD}$ are strictly convex quadratic programs. The constraints of LQP consist of the nonnegativity conditions $\lambda \geq 0$, whereas $\overline{GD}$ also includes the single equality constraint $(b^T+q^TQ^{-1}A^T)\lambda = 1$. The following theorem shows the relationship between the gauge dual programs GD or $\overline{GD}$ and the Lagrange dual LQP:

Theorem 3. If Q is positive definite, then

(i) If λ^* is a solution to GD and $t^* = \lambda^{*T}AQ^{-1}A^T\lambda^*$, then

a) $t^* \neq 0$ if and only if $\bar{\lambda} = \lambda^*/t^*$ solves the Lagrange dual LQP,

b) $t^* = 0$ if and only if QP is infeasible.

(ii) If $\bar{\lambda}$ is a solution to LQP and $\bar{t} = b^T\bar{\lambda} + q^TQ^{-1}A^T\bar{\lambda}$, then

a) $\bar{t} \neq 0$ if and only if $\lambda^* = \bar{\lambda}/\bar{t}$ solves the guage dual GD,

b) $\bar{t} = 0$ if and only if QP has a solution $\bar{x}$ to $A\bar{x} \geq b$ satisfying $M\bar{x} + s = 0$, i.e., if and only if GD is infeasible.

PROOF: The proof follows from an examination of the Karush-Kuhn-Tucker conditions for GD and LQP. The transformations follow from direct substitution. [X]

Q is Positive Semi-Definite

We now turn our attention to the broader case, when Q is positive semi-definite and q lies in the row space of M, whence $q = M^Ts$ for some vector $s \in R^n$. In this case, the Lagrange dual of QP is

$$\begin{array}{lll} \underset{\lambda,x}{\text{maximize}} & b^T\lambda - (1/2)x^TM^TMx & \\ \text{subject to} & A^T\lambda - M^TMx = M^Ts & (LQP'), \\ & \lambda \geq 0 & \end{array}$$

as in Dorn [1]. The gauge dual of GP is

$$\begin{array}{lll} \underset{\lambda,u}{\text{minimize}} & \|u\|_2 & \\ \text{subject to} & -A^T\lambda + M^Tu = 0 & (GD'), \\ & b^T\lambda + s^Tu = 1 & \\ & \lambda \geq 0 & \end{array}$$

which is equivalent to the quadratic program

$$\begin{array}{lll} \underset{\lambda,u}{\text{minimize}} & u^Tu & \\ \text{subject to} & -A^T\lambda + M^Tu = 0 & (\overline{GD}') \\ & b^T\lambda + s^Tu = 1 & \\ & \lambda \geq 0 & \end{array}$$

Analogous to Theorem 3, Theorem 4 demonstrates the relationship between the two different dual quadratic programs LQP' and $\overline{GD}'$.

Theorem 4. If Q is positive semi-definite, $Q = M^TM$ and $q = M^Ts$ for some $s \in R^n$, then

(i) (λ^*, u^*) constitute an optimal solution to the gauge dual $\overline{GD}'$ if and only if there exists x^*, t^* such that

(a) $-A^T\lambda^* + M^Tu^* = 0$

(b) $b^T\lambda^* + s^Tu^* = 1$

(c) $\lambda^* \geq 0$

(d) $Ax^* \geq bt^*$

(e) $u^* = Mx^* + st^*$

(f) $\lambda^{*T}Ax^* = \lambda^{*T}bt^*$

(ii) $t^* = 0$ if and only if QP is infeasible. $t^* \neq 0$ if and only if $\bar{\lambda} = \lambda^*/t^*$, $\bar{x} = x^*/t^*$ constitute a solution to the Lagrange dual LQP'.

(iii) $(\bar{\lambda},\bar{x})$ constitute an optimal solution to the Lagrange dual LQP' if and only if there exists $\bar{z}$ such that

(a) $\bar{\lambda} \geq 0$

(b) $A^T\bar{\lambda}-M^TM\bar{x}=M^Ts$

(c) $A\bar{z} \geq b$

(d) $M^TM\bar{x}=M^TM\bar{z}$

(e) $\bar{\lambda}^TA\bar{z}=\bar{\lambda}^Tb$

(iv) $\bar{\lambda}^Tb+s^Ts+s^TM\bar{z}=0$ if and only if QP has a solution $Ax \geq b$, $Mx+s=0$, i.e. if and only if $\overline{GD}'$ is infeasible.

$\bar{\lambda}^Tb+s^Ts+s^TM\bar{z}\neq 0$ if and only if (λ^*,u^*) solves the gauge dual $\overline{GD}'$ with multipliers x^*, t^* given by:

$t^* = 1/(\bar{\lambda}^Tb+s^Ts+s^TM\bar{z})$

$\lambda^* = \bar{\lambda}t^*$

$u^* = (M\bar{z}+s)t^*$

$x^* = \bar{z}t^*$. [X]

The conditions (i) and (iii) of this theorem are simply the Karush-Kuhn-Tucker conditions, and the transformations in (ii) and (iv) follow from direct substitution.

B. Programs with the l_p Norm

If $f(x) = \|x\|_p$, $1 \leq p \leq \infty$. then the polar of $f(\cdot)$ is $f^o(y) = \|y\|_q$, where q must satisfy $1/p+1/q = 1$. In particular, $p=1$ or $p=\infty$ if and only if $q=\infty$ or $q=1$, respectively.

Consider the l_p norm program:

$$\begin{aligned} &\underset{x,t}{\text{minimize}} && \|w\|_p \\ &\text{subject to} && Bw+Cz \geq d \end{aligned} \qquad (GP_p)$$

The gauge dual of GP_p, derived using $\bar{P}$ and $\bar{D}$, is

$$\begin{array}{lll} \underset{\lambda}{\text{minimize}} & \|B^T\lambda\|_q & \\ \text{subject to} & C^T\lambda = 0 & \\ & d^T\lambda = 1 & (GD_q) \\ & \lambda \geq 0 & \end{array}$$

Because $f(x) = \|x\|_p$ satisfies the projection qualification, the results of Theorem 2 are valid for GP_p and GD_q.

Note that the program GP, which is a derived equivalent of the quadratic program QP (when Q is positive semi-definite and $q=M^Ts$ has a solution s), can be cast as an instance of GP_p, by setting

$$B = \begin{bmatrix} I \\ -I \\ 0 \end{bmatrix}, \; C = \begin{bmatrix} -M \\ M \\ A \end{bmatrix}, \; d = \begin{bmatrix} s \\ -s \\ b \end{bmatrix}, \text{ and } p = 2.$$

Thus GP_p is a more general program than QP.

When $p \neq 1$, $p \neq \infty$, the program GP_p is equivalent to

$$\begin{array}{lll} \underset{x,z}{\text{minimize}} & (1/p)\ \|w\|_p^p & \\ \text{subject to} & Bw+Cz \geq d & (LP_p) \end{array}$$

The Lagrange dual of this program is

$$\begin{array}{ll} \text{maximize } \lambda^T d-(1/q)\|B^T\lambda\|_q^q & \\ \text{subject to } C^T\lambda = 0 & (LD_q) \\ \qquad\qquad\quad \lambda \geq 0 & \end{array}$$

The gauge dual GD_q and the Lagrange dual LD_q bear a relationship that generalizes the case of quadratic programming in Theorems 3 and 4. This relationship is demonstrated below. In the theorem, the notation x^p, where $x \in R^n$, denotes the vector whose j^{th} component is $(\text{sign } x_j)(|x_j|^p)$.

Theorem 5. If $p \neq 1$ and $p \neq \infty$, then

(i) If λ^* is an optimal solution to the gauge dual GD_q and $t^* = \lambda^{*T}(B)(B^T\lambda^*)^{q-1}$, then

(a) $t^* \neq 0$ if and only if $\bar{\lambda}=\lambda^*/t^*$ solves the Lagrange dual LD_q.

(b) $t^*=0$ if and only if GP_p is infeasible.

(ii) If $\bar{\lambda}$ is an optimal solution to the Lagrange dual LD_q, then

(a) $\bar{\lambda}^T d \neq 0$ if and only if $\lambda^*=\bar{\lambda}/\bar{\lambda}^T d$ solves the gauge dual GD_q.

(b) $\bar{\lambda}^T d=0$ if and only if GP_p has an optimal solution with value 0, i.e. $Cz \geq d$ has a solution. [X]

The proof of this theorem follows from examining the Karush-Kuhn-Tucker conditions and substituting in the transformations as given.

Although the programs GP_p and LP_p are equivalent (their objective functions differ by a monotone transformation), the gauge dual GD_q of GP_p will yield a better (i.e., larger) lower bound on the optimal solution to GP_p than will the Lagrange dual LD_q for LP_p. To see this, let (w,z) be any strongly feasible solution to GP_p and LP_p, and let ϕ and ψ be the corresponding objective values of (w,z) in GP_p and LP_p, namely $\phi = \|w\|_p$, $\psi = (1/p)\ \|w\|_p^p$. Let $\lambda \geq 0$ be any feasible solution to LD_q with a positive objective value, and hence $\bar{\lambda} = \lambda/(d^T\lambda)$ is a feasible solution to GD_q. Let $g = \|B^T\bar{\lambda}\|_q$, and $h = d^T\lambda - (1/q)\|B^T\lambda\|_p^p$ be the corresponding objective function values of $\bar{\lambda}$ and λ in the programs GD_q and LD_q, respectively. Then 1/g and h each represent a lower bound on the optimal primal objective values for GP_p and LP_p, and the values ϕg and ψ/h are numbers greater than or equal to one that measure the duality gap in the respective dual pairs of programs, as a ratio of the primal objective function value to the dual objective function value. We have:

Lemma 3. Under the above assumptions, $\phi g \leq \psi/h$.

This lemma states that the corresponding duality gaps (measured as a ratio) is always smaller for the gauge dual GD_q than for the Lagrange dual LD_q. Of course, the comparison is somewhat unfair, inasmuch as the objective functions of GP_p and LP_p differ by a monotone transformation. Yet the proof below shows that the gauge dual GD_q in a sense uses an intrinsically better convex inequality than does the Lagrange dual LD_q.

Proof of Lemma 3: Let w, z, ϕ, ψ, g, h, λ, and $\bar{\lambda}$ be as stated. Then $\lambda^T d \leq \lambda^T Bw + \lambda^T Cz = \lambda^T Bw \leq \|w\|_p \|B^T\lambda\|_q \leq (1/p)\|w\|_p^p + (1/q)\|B^T\lambda\|_q^q$,

the last inequality being an instance of the inequality between the arithmetic and the geometric mean. Let s, t, and u represent the nonnegative gaps in the three inequalities above, respectively.

Then we have

$$\phi g = \|w\|_p \|B^T\lambda\|_q/(d^t\lambda) = \frac{d^T\lambda+s+t+u}{d^T\lambda} = 1 + \frac{s+t+u}{d^T\lambda}$$

Therefore

$$\phi g = 1 + \frac{s+t+u}{d^T\lambda} \leq 1 + \frac{s+t+u}{d^T\lambda-(1/q)\|B^T\lambda\|_q^q} = \frac{\psi}{h} . \quad [X]$$

Note that it is the inequality between the arithmetic mean and the geometric mean that drives the result.

C. Linear Programming

As a final note, observe that the linear programming problem:

$$\text{LP:} \qquad \text{minimize} \quad c^T x$$
$$\text{subject to} \quad Ax \geq b$$

can be formulated as a gauge program when the optimal value of LP is positive. In this case, let $f(x) = \max \{c^T x, 0\}$. Then $f(\cdot)$ is a gauge and $C = \{x \in R^n | f(x) \leq 1\} = \{x \in R^n | c^T x \leq 1\}$. It is then straightforward to compute $C^o = \{y \in R^n | y = cv, \ v \geq 0\}$, and

$$f^o(y) = \begin{cases} v & \text{if } y = cv \text{ for some } v \geq 0 \\ +\infty & \text{else} \end{cases} \quad .$$

The gauge dual of LP then is

$$\text{DLP}': \qquad \underset{\lambda, v}{\text{minimize}} \quad v$$
$$A^T\lambda - cv = 0$$
$$b^T\lambda = 1$$
$$\lambda \geq 0$$

which is equivalent to the standard linear program dual:

$$\text{DLP:} \qquad \underset{\lambda}{\text{maximize}} \quad b^T\lambda$$
$$\text{subject to} \quad A^T\lambda = c$$
$$\lambda \geq 0 \quad .$$

Acknowledgment

I am very grateful to Professor Thomas Magnanti for his valuable comments and careful reading of a draft of this paper.

Appendix: Separation Theorems

Given two convex sets X and Y in R^n, X is properly separated from Y by a hyperplane H provided X and Y lie in the opposite closed halfspaces bounded by H, and X and Y do not both lie in H. The following theorem of Fenchel [3] characterizes when X and Y can be properly separated:

Theorem 6 (Fenchel [3], see Rockafellar [12]). Let X and Y be nonempty convex sets in R^n. X and Y can be properly separated by a hyperplane H if and only if (rel int X) $\cap$ (rel int Y) = ϕ. [X]

If X and Y are convex sets in R^n, X is openly separated from Y by a hyperplane H provided X lies in one of the open halfspaces bounded by H and Y lies in the other closed halfspace. Clearly, if X is openly separated from Y, then X and Y are properly separated. Klee's results of [8] give criteria on X and Y that are sufficient for X to be openly separated from Y by some hyperplane H. Some of these criteria are given below.

A convex set $X \subset R^n$ is called evenly convex [4] provided that X is the intersection of a family of open halfspaces. A set $Z \subset R^n$ is called an asymptote of a convex set $Y \subset R^n$ provided that Z is an affine variety, $Z \cap Y = \phi$, and $\inf \{\|z-y\| \mid z \in Z, y \in Y\} = 0$. The set $Y \subset R^n$ is said to be boundedly polyhedral if its intersection with any bounded polyhedron is a bounded polyhedron. One of Klee's results in [8] is the following:

Theorem 7 (Klee [8]) If X and Y are disjoint convex subsets of R^n, then X can be openly separated from Y if X's projections are all evenly convex, and Y admits no asymptote and Y is boundedly polyhedral. [X]

Actually, Klee's results are much broader than indicated. He shows that the stated conditions are maximal in a sense he defines precisely, and he also gives five alternative criteria that guarantee that X can be openly separated from Y.

Remark 5. If X and Y are disjoint convex sets and all projections of X are closed, and Y is a polyhedron, then there exists $y \in R^n$ and $\alpha \in R$ such that $y^T x < \alpha$ for all $x \in X$ and $y^T x \geq \alpha$ for all $x \in Y$.

PROOF: Any closed convex set is evenly convex, since any closed convex set is equal to the intersection of the family of the closed halfspaces that contain it, see Rockafellar [12], and each closed halfspace is the intersection of an infinite family of open halfspaces. Thus, if all projections of X are closed, all projections of X are evenly convex. If Y is a polyhedron, then Y is boundedly polyhedral, and admits no asymptote. Thus X and Y satisfy the conditions of Theorem 7, whereby the desired result is obtained. [X].

References

[1] W.S. Dorn, "Duality in quadratic programming," Quart. Appl. Math. 18, 155-162, 1960.

[2] E. Eisenberg, "Duality in homogeneous programming," Proceedings of the American Mathematical Society 12, 783-787 (1961).

[3] W. Fenchel, "Convex cones, sets, and functions," lecture notes, Princeton University, 1951.

[4] W. Fenchel, "A remark on convex sets and polarity," Comm. Sem. Math. Univ. Lund. (Medd. Lunds Univ. Math. Sem.) Tome Supplementaire, 1952, 88-89.

[5] D.R. Fulkerson, "Blocking and anti-blocking pairs of polyhedra," Mathematical Programming 1, 168-194 (1971).

[6] C.R. Glassey, "Explicit duality for convex homogeneous programs," Mathematical Programming 10, 176-191 (1976).

[7] J. Gwinner, "An Extension lemma and homogeneous programming," J. Opt. Theory and Applications 47, 321-336 (1985).

[8] V. Klee, "Maximal separation theorems for convex sets," Transactions of the American Mathematical Society, 134 1, 1968.

[9] D.G. Luenberger, Optimization by vector space methods, John Wiley & Sons (New York, 1969).

[10] T.L. Magnanti, "Fenchel and Lagrange duality are equivalent," Mathematical Programming 7, 253-258 (1974).

[11] L. McLinden, "Symmetric duality for structured convex programs," Trans. Amer. Math. Soc. 245, 147-181 (1978).

[12] R.T. Rockafellar, Convex analysis, Princeton University Press (Princeton, New Jersey, 1970).

[13] P.H. Ruys and H.M. Weddepohl, "Economic theory and duality," in: J. Kriens, ed., Convex analysis and mathematical economics, Tilburg 1978, Springer (Berlin, 1979).